Learning Resource

Demonstrate knowledge of industrial process control

Level 4 | Credits 2

Te Pūkenga

Contents

Using this resource

The following information boxes may be found in this resource.

If there is an important note, or a key idea – they may be written into a box like this.

Definition

Word Meaning of word

$f(x)$

Formula

If there is an important formula, it may appear in a box like this.

$I = \sqrt{I} \div R$

Part 1: Common transducers

A **transducer** is a device that converts energy from one form to another. A transducer is therefore the first element in a measuring system, which takes information about any variable being measured and transforms this into a more suitable or appropriate quantity for the attached control system.

This is done by sensing relative changes in the characteristics of various materials and interpreting and converting the changes into electrical signals. These characteristics include:

- Resistance
- Capacitance
- Voltage
- Physical dimensions
- Pressure
- Relative position
- Light intensity.

The changes are detected by sensors that form part of the transducer.

For the purposes of our study, a **transducer** can be re-defined as a device that converts physical quantities (heat, light, pressure, rotation) into electrical signals.

Transducers may have one or more parts. A simple transducer may contain one part, such as a thermocouple, that directly converts heat into electrical energy. A more complicated transducer may contain a sensor together with some electronic circuitry, such as a microphone with piezoelectric crystal and amplifier. A transducer which produces a 0-20mA or 4-20mA output signal is sometimes referred to as a transmitter.

As an electrical worker, you need to understand the functions, principles of operation and applications of different types of transducers. Some examples are:

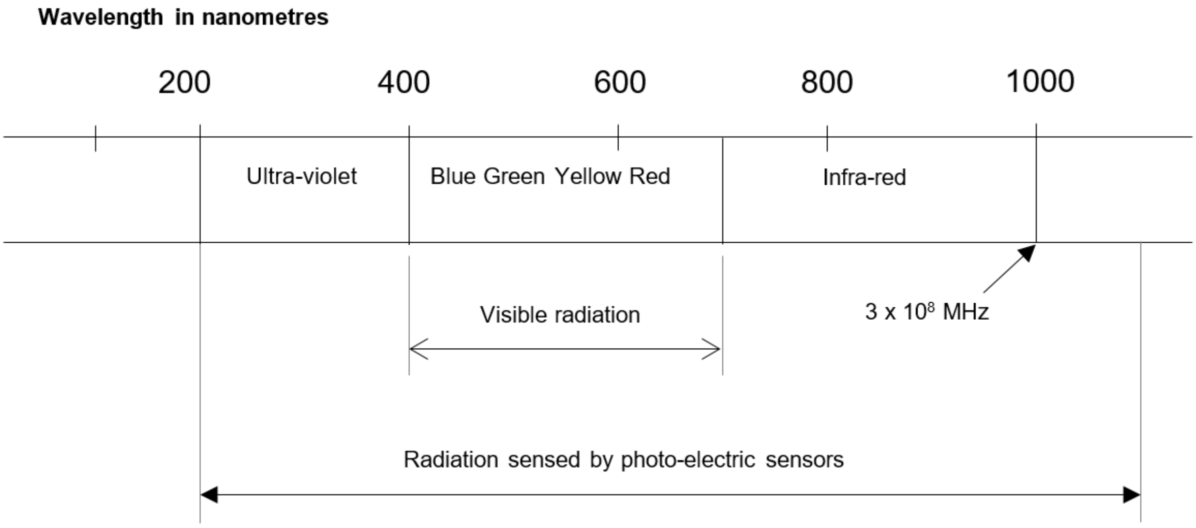
- Light sensing transducers
- Humidity sensors
- Temperature transducers
- Pressure and strain transducers
- Level and proximity transducers,

Light sensing transducers

Light transducers capture light intensity and convert it into an electrical signal. Some common uses of the light transducer are in production rooms where it is necessary to maintain the same level of brightness or in places where light is required to activate when daylight intensity decreases (for example, street lights).

The light spectrum used by the transducer can fall beyond the visible range and include ultra-violet and infra-red.

The diagram below, shows the light spectrum and range used by photoelectric sensors.



Illuminance is the intensity of the light that falls on a surface. It is measured in lux (lx). The luminous sensitivity of a light transducer is the amount of current output you get for a particular intensity of light.

For example, you may get 200µA when the device is exposed to 50 lx meaning that the sensitivity is 40 µA / lx.

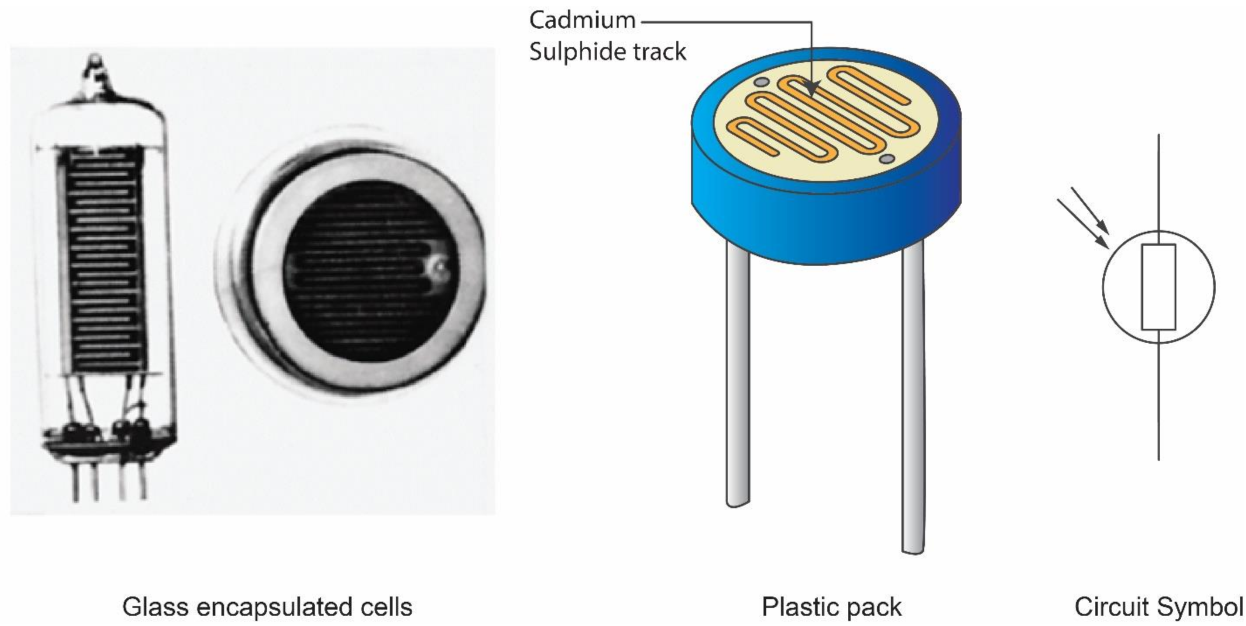
The following section looks at different types of light transducers.

Semiconductor Light Transducers

Light Dependant Resistor

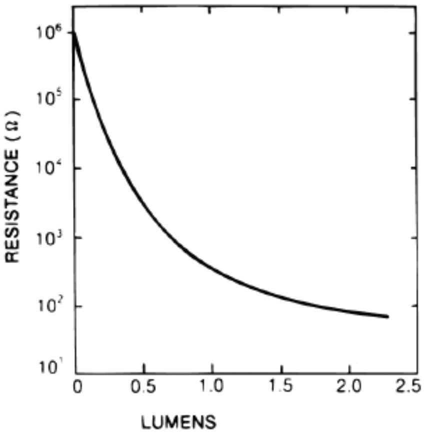
The light-dependent resistor (LDR) is a photoconductive device. Light falling on the sensitive surface causes the device to become more conductive and its resistance to fall. A common LDR is made from *cadmium sulphide*, also called a CdS cell.

The LDR or CdS cell consists of two metal grids with CdS in between the metal grids. The cell is usually enclosed in a glass envelope to exclude dirt and moisture. Modern versions are more often encased in clear plastic.



The LDR

The light dependant resistor cell measures and responds to light in the visible range, with maximum response to red light. The resistance of the cell will change when light falls on it. In darkness, its resistance can be many megohms. The graph (right), shows the response characteristic of an LDR. The resistance falls to a much lower value when more light falls on the cell face.

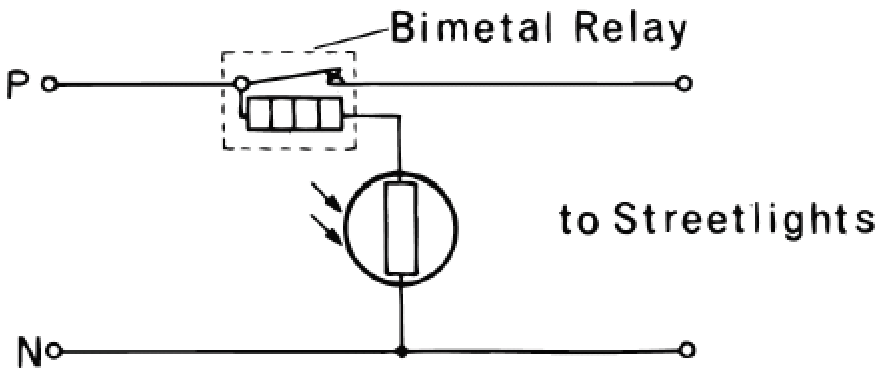


Response characteristic

These cells are mainly used for on/off switching. They can be used with alternating or direct current supplies and may dissipate up to 0.5 watts. They are very sensitive, giving about 180 μA / lx luminous sensitivity, but they respond very slowly to changes of light.

Because of its slow response, the LDR cannot be used in high-speed counting. It may be used in flame failure circuits for boilers, level control, or as a day/night switch for twilight switching of external flood lights.

The diagram below, is a diagram of a **typical twilight switch**.



Twilight switch

During daylight, the LDR is at low resistance and allows heater current to flow. This heats the bimetal strip which expands and remains bent, holding the contacts open.

When daylight fades, the bimetal strip cools down and contracts, thus allowing contact to be made resulting in the light being switched on.

- The response of the bimetal is slow, so that temporary changes of light do not cause the switch to close.
- Twilight switches are reliable and used to be very popular. The drawback is that they always draw current and the switching point is subject to ambient temperature changes.
- These devices have therefore generally been replaced by modern electronics, which still use a CdS cell as the sensor but use thyristors to do the actual switching.

A LDR should not be exposed to strong light such as direct sunlight, nor should it be exposed to extremes of heat.

The silicon photodiode

The silicon photodiode is a reverse-biased silicon diode with provision made to allow light to fall on the p-n junction.

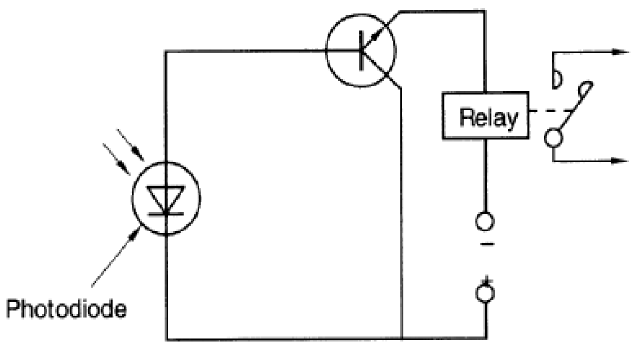
Current flow in the reverse-biased diode is very small and is usually called **leakage current**. In the photodiode, this leakage current, in the absence of light, is called **dark current**. The dark current is typically about 2 μA and represents a resistance of several megohms.

Silicon devices also respond to infrared light. When light falls on the junction, covalent bonds are broken, causing the reverse resistance to fall. When light falls on the junction, the apparent resistance falls, allowing the reverse current to rise, from a few microamperes when dark to about 30 μA at 1000 lx. Luminous sensitivity ranges between 0.03 and 0.15 mA/lx, which is much less than the sensitivity of the LDR. However, the silicon photodiode is much faster acting.

As the photodiode can only pass microamperes it is not practical on its own. We need to 'amplify' this output current to use the device in a practical application.

The diagram (right), shows a simplified circuit for a daylight switch.

- An n-p-n transistor is used as a common collector amplifier to feed a relay.
- The transistor receives its forward bias from the reverse-biased photodiode.
- When sufficient light is focused onto the diode junction, it will allow about 20 μA to flow.



Daylight switch

- This current flows from the supply positive, through the photodiode, through the base emitter junction of the transistor, and through the relay coil and back to negative.

The transistor turns on and the relay energises. Removal of the light causes the diode current to fall to a few μA, and the relay de-energises.

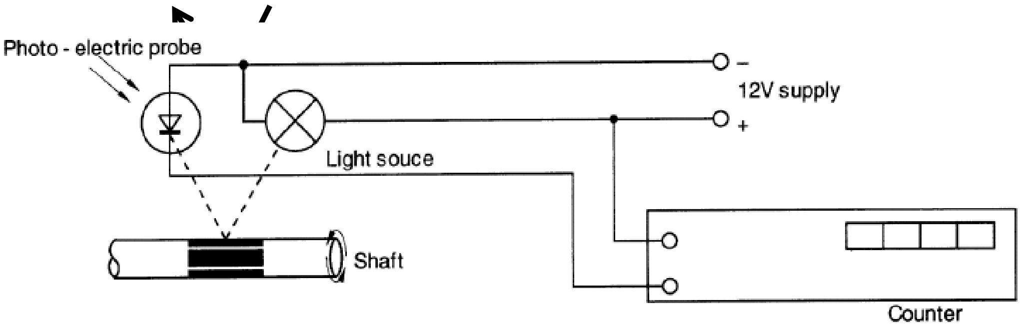
Photo Diode Tachometer

The response time of a silicon photodiode is typically about 0.01 μs, but response times of 0.001 μs are possible. These diodes may be used in such light-beam-interruption high-speed counting applications as electronic tachometers.

The diagram below shows the principle of a reflective tachometer, where the light is aimed at a rotating shaft.

The shaft is marked with dark and light stripes, so that light is reflected from the stripes on the shaft to the diode. A counter counts the pulses and displays shaft speed in revolutions per minute.

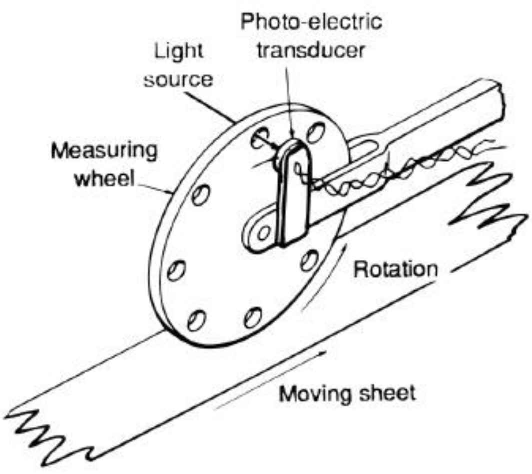
This tachometer is also called a **digital speed transducer**.



A photodiode tachometer

The diagram below, shows a greatly simplified drawing of another form of digital speed transducer used to measure linear speed, that is, speed in a straight-line. The transducer wheel is in contact with a moving sheet of metal, whose motion and speed are to be measured.

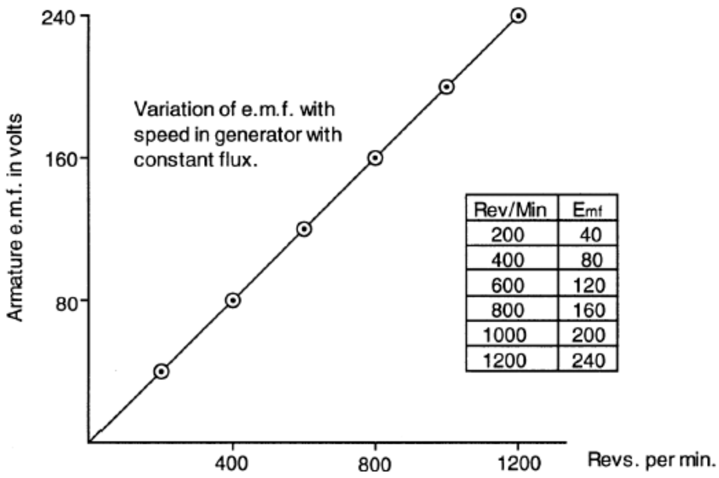
The photoelectric transducer picks up light passing through the holes in the wheel. Each revolution of the wheel produces eight light pulses, which can be converted to speed of motion or length. This pulse disc device can be used to measure the speed of a conveyor belt or a production process, such as steel strap being fed into a cutting machine (e.g. a cut to length application).



Measuring linear speed

Analogue speed transducers (tacho-generators)

The photodiode is used as a digital speed transducer, where voltage pulses are produced at a frequency that depends on the speed of a rotating shaft or wheel.



Graph of speed of armature and generated e.m.f.

In the analogue speed transducer, a voltage is produced that is proportional to speed. A small generator is turned by the rotating part whose speed is to be measured.

Separately excited generators (or self-exciting generators), such as those with permanent magnet fields that have a constant field flux, generate a voltage that is proportional to speed.

The above diagram shows a graph of generated voltage against armature speed. For the separately excited generator, this graph will be a straight line.

The output of the generator is connected to a voltmeter that has its scale calibrated in revolutions per minute instead of volts, as shown in the diagram (right).

Analogue speed indicator

The Silicon Phototransistor

The silicon phototransistor is a bipolar junction transistor (BJT) with a transparent lens that focuses light onto the thin base region.

Under **dark** conditions, the dark current (leakage) is usually less than 1 μA.

When **light** is applied to the transistor:

- The breaking of covalent bonds provides base current.
- Amplification takes place, and so the phototransistor is much more sensitive than the photodiode.
- Luminous sensitivity is typically, 10 μA/lx, which is much less than that of the LDR.

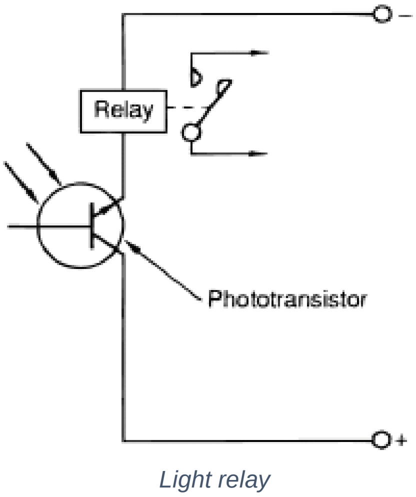
A phototransistor has a response time of about 5 μs and so is much slower than the photodiode. However, this response is fast enough for these transistors to be used in counters, optical couplers, and many on–off operations that require a relatively fast, consistent response.

The diagram below, shows a phototransistor equivalent of the photodiode circuit. The phototransistor senses the light and energises the relay.

A practical circuit would include capacitors or a freewheeling diode across the relay coil. These components protect the transistor from high voltage produced by the inductance of the coil.

For transistor action to occur in a phototransistor, the junctions must be biased correctly. Thus, the collector-to-emitter supply must be **direct current**.

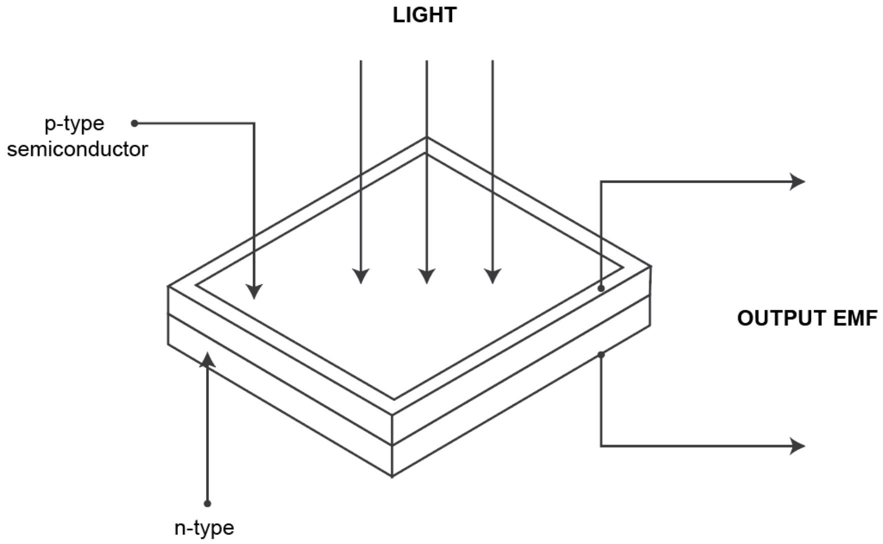
Some phototransistors do not have a base connection. Those that do have a base connection may be biased from the collector supply, as is usual with the BJT.



Light relay

Solar cell

A solar cell is a device that converts light into electricity. It is also known as a **photovoltaic (PV) cell**. It is made from specially treated silicon.



Solar Cell

Commercial cells produce about 0.4V at about 1.5A.

When light strikes the cell, a certain portion of it is absorbed within the semiconductor material and the energy of the absorbed light is transferred to the semiconductor. The energy 'knocks' electrons loose, allowing them to flow freely.

Within the PV cells are one or more electric fields that act to force electrons freed by light absorption to flow in a certain direction.

By placing metal contacts on the top and bottom of the PV cell, that current can be drawn off to use externally.

When connected in series to form a bank of cells, PV cells are frequently used as a power supply for remote instruments such as weather stations. Another example would be on a solar-powered calculator.

The output of a single cell is relatively small, and so it can be used to sense and measure light and operate as a transducer to drive a process control system.

Resistive humidity sensors

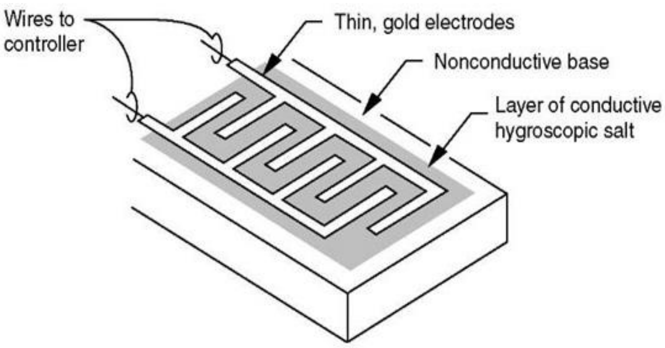
Some materials are **hygroscopic**, that is, they absorb water from the atmosphere. This results in the following features of the material undergoing change:

- Resistivity
- Permittivity
- Mass

Resistive humidity sensors

The effects of these changes can be applied to several types of sensors to obtain an electrical output. The most common of these sensors is the resistive humidity sensor.

Metallic salts such as lithium chloride are often used as the sensitive material in electrolytic resistive humidity sensors because they show a large change in resistivity for a small change in moisture content. The salts are deposited between intermeshing printed-circuit electrodes on a rigid substrate.



Resistive humidity sensors arrangement

Resistive sensors are based on an inter-digitated or bifilar winding. The conductors are arranged as in the example (left), to provide maximum area of contact with the sensitive material. After deposition of a hydroscopic polymer coating, their resistance changes inversely with humidity.

The air or gas to be tested flows past the sensor, so that the moisture content can affect the sensor resistivity. If the humidity increases, the resistance decreases, because more moisture is absorbed.

If the air becomes drier, the sensor gives back some of the absorbed water, and the resistance increases.

The main disadvantages of the resistive humidity sensor are:

- Slowness of response, which can take up to several minutes
- Sensor’s sensitivity to temperature.

Because the cell uses an electrolyte as the sensitive material, a DC voltage cannot be used to determine the resistance. The frequency of the AC supply should be above 20 Hertz, and the voltage must be kept low to avoid cell breakdown.

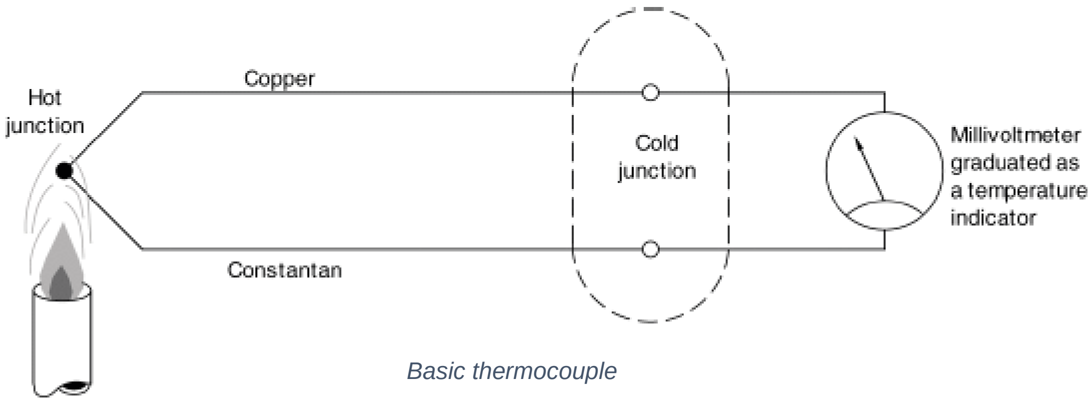
A typical maximum current for a cell is 1mA, with a resistance change from 35,300 KΩ at 5% relative humidity to 80 KΩ k at 100% relative humidity.

The Thermocouple

A thermocouple is a **thermoelectric transducer** that produces its own EMF when heated and can be used to measure temperature. The thermocouple consists of two dissimilar conductors (such as copper and constantan), welded together, as shown in the diagram below. (Constantan is an alloy made from 45% nickel and 55% copper).

A copper-constantan thermocouple is used over a recommended temperature range of –300 °C to +400 °C. The average thermal emf is 0.041 mV/°C (41 μV/°C).

- That is, for every degree rise in temperature, there is 0.041 mV rise in the EMF produced. This relationship is very linear and therefore very predictable.



Basic thermocouple

Another suitable thermocouple combination is an iron-constantan, which operates in the temperature range –210 °C to +1200 °C. The average thermal emf is 0.055 mV/°C (55 μV/°C).

Other combinations allow temperatures of up to 1700 °C to be measured.

Operation of the thermocouple

> An electromotive force (EMF) is created whenever there is a temperature difference within a conductor. The EMF produced is proportional to the temperature difference between the cold and hot ends of the conductor.

This EMF is different for different materials. When the two different materials are joined at one end and heated, there will be difference of potential between the two cold ends. This EMF can be measured and will be proportional to the temperature difference between the hot and cold ends. The small physical size and the wide temperature range of the thermocouple make it a most versatile transducer.

Thermocouples are used for temperature measurement in many process control and alarm systems. A basic system could measure the temperature in ovens and kilns. The main disadvantage of the thermocouple is its low sensitivity — typically, about 5 mV for 100 °C difference between the hot and cold junctions.

Ranges and Types:

Most instrument suppliers provide look up tables for thermocouples which feature their safe operating and measuring range, e.g. type K Ni-Cr (Chromel) or Ni-Al (Alumel) -330 to 1370 °C whereas a typical J type Iron / Cu-Ni (Constantan) - 210 to 1200 °C.

Compensating cable:

Measurement is often taken remotely from the thermocouple cold ends, and in this case the thermocouple cable needs to be extended.

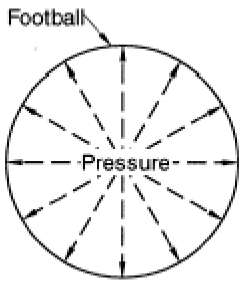
Cheaper compensating cable can be used to extend the thermocouple leads. Matching compensating cable, with similar thermoelectric properties to the thermocouple cable, is used which ensures the integrity of the millivolt signal over its length to the measuring or recording instrument.

Instrument suppliers provide look-up charts for the temperature range, thermocouple joint and the appropriate compensating cable.

Pressure and Strain Transducers

Pressure

Pressure is the force exerted by a liquid or gas on the walls of its container, in opposition to being pressed, or confined. For example, the air in a football presses against its casing.



In industry, the pressure of such things as gas, hydraulic fluid, steam or molten plastics is measured. The unit of pressure is the **pascal** (Pa), which is equal to a force of one newton per square metre.

Strain

Strain is the effect of **changing the shape of a solid** when a force is applied.

Strain is expressed as a ratio of the change in length over the original length.

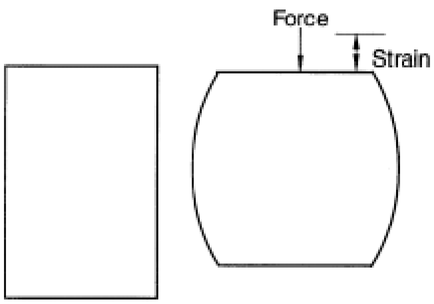
$f(x)$	**Formula** Strain = (Contraction or extension) / (original length)

Stress

Stress is the **force per unit area**, measured in pascals that causes strain.

Example (a) shows a solid column standing freely. When a force is applied to the top of the column, as shown in the example (b), the column changes shape, resulting in strain.

This can be observed from its reduced height and greater width. You can see the effects of strain for yourself when you apply force across the ends of a pencil eraser.



(a) *(b)*

Stress and train

Resistive Strain Gauge

A resistive strain gauge is based on the principle that resistance changes when strain results from a force applied to resistance wire. This principle can be used to measure movement.

Consider a piece of resistance wire held taut between two posts. If you strain the wire by moving the posts further apart, the following occurs:

- The length of the resistance wire increases, and its diameter decreases. Both changes cause the resistance to increase.
- The stress on the atomic structure within the wire also causes the resistance to increase.

The resistive strain gauge using a taut resistance wire is relatively large. It must always be held in tension, and it can measure only extension. These problems are overcome in the bonded strain gauge.

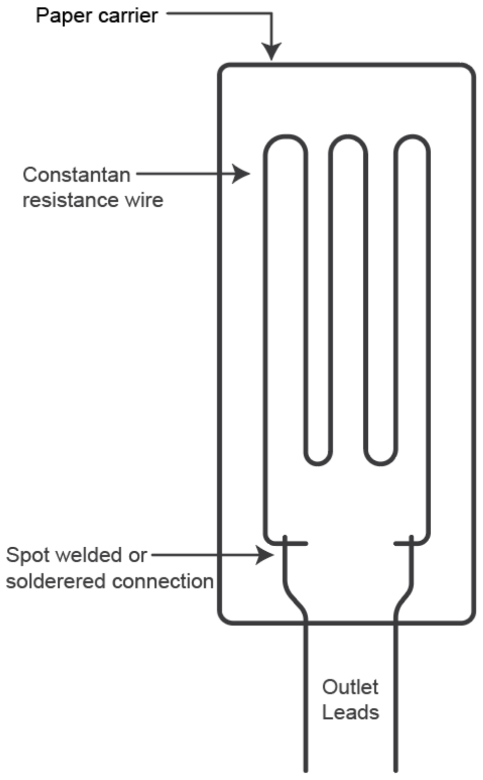
The bonded resistive strain gauge consists of several small loops of resistance wire bonded to a thin backing sheet.

The thin backing, or base carrier, is made from *nitrocellulose*, impregnated paper or plastic and impregnated fibreglass.

The resistance wire is made from alloys of copper, manganese and aluminium, copper, manganese and nickel, or copper and nickel (constantan).

When the strain gauge is firmly bonded to the sample under test, the gauge becomes part of the sample. If the sample is placed under compression, the strain gauge also experiences the compression.

Similarly, when a sample experiences stretching, the gauge also stretches.

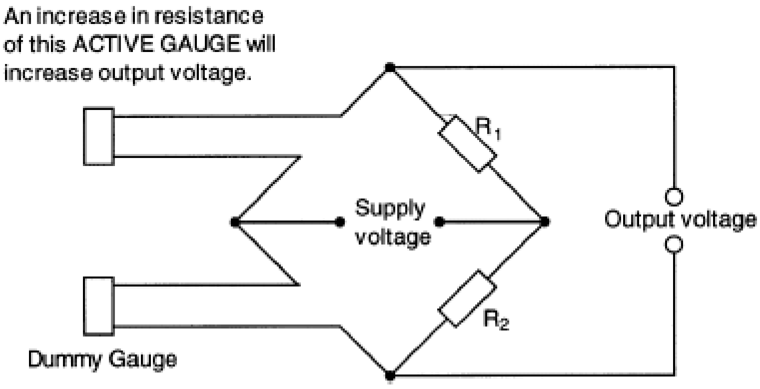


Bonded strain gauge

In this way, an increase in gauge resistance indicates extension, and a decrease in gauge resistance indicates compression.

Measuring Circuit

To overcome temperature effects, two bonded strain gauges are connected, as in the diagram below. Only one gauge is bonded to the piece under test. The second gauge is held in contact with the test piece. In this way, both gauges experience the same temperature effects, which are then cancelled out in the bridge.

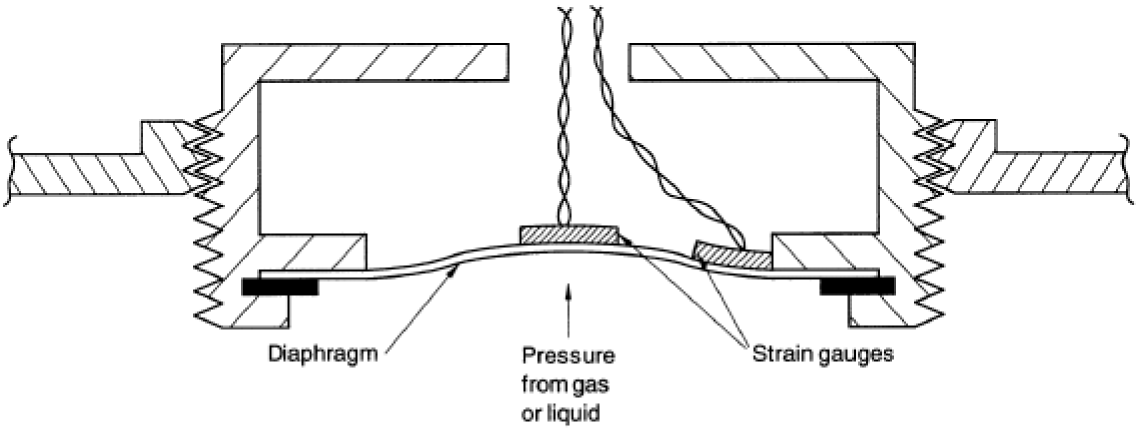


Strain gauge bridge circuit

This property of the gauge can be used to detect movement in a building can be used to detect movement in a building. A gauge for instance, placed on a crack in a wall, can detect small movements in the building before the situation gets too bad.

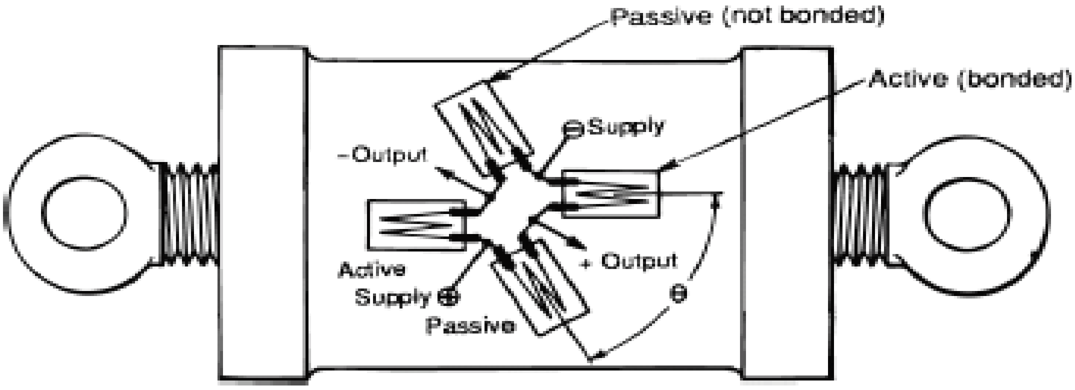
The supply to the bridge may be alternating current or direct current, depending on the kind of output that is required. An amplifier is usually needed, because the bridge output is small.

Strain gauges have a wide range of applications. The example below shows their use in a pressure transducer.



Pressure transducer

Both strain gauges are bonded to the diaphragm. The centre strain gauge experiences a stretching strain, and the outer gauge experiences a compression strain. Both the gauges are active because they experience opposite kinds of strain. This arrangement doubles the effect of the strain as compared to a single active strain gauge and cancels the effects of temperature changes.



Load cell

The diagram above shows a load cell, placed between a crane and a load being lifted, to measure the weight of the load. A bridge circuit is used with a strain gauge in all four limbs. The active gauges are strained by the load and produce an output to indicate the weight.

The two passive strain gauges are required in this application to complete the bridge and for temperature compensation.

Bonded strain gauges may be used to measure strain that occurs quite slowly. A bonded strain gauge can also be used to measure rapidly changing strains.

Frequencies of up to 50 KHz can be measured with such gauges.

The latest kind of **piezoresistive** strain gauge uses semiconductor material to produce a much greater resistance change for a given strain.

Therefore, this kind of strain gauge is much more sensitive to small changes in pressure.

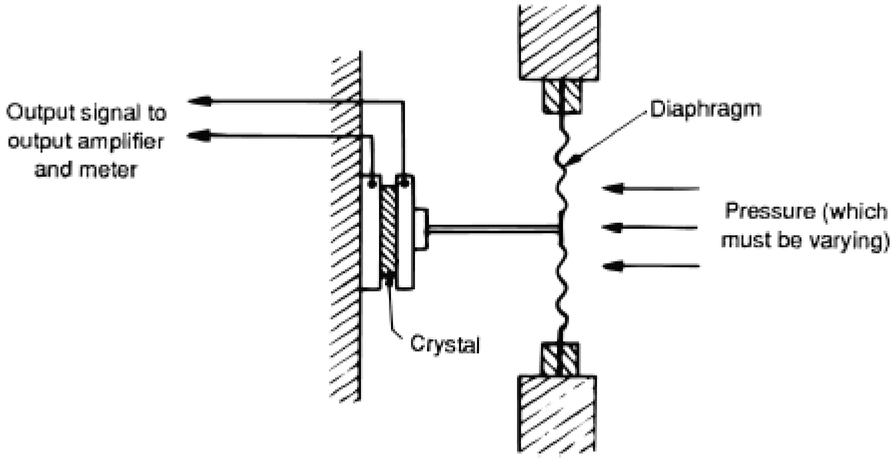
The Piezoelectric Transducer

The **Piezoelectric Effect** is the ability of certain materials to generate an electric charge in response to applied mechanical stress. The word *piezoelectric* is derived from the Greek *piezein*, which means to squeeze or press, and piezo, which is Greek for “push”.

Some crystalline materials produce an EMF when squeezed or twisted. The most common materials used are ceramics, *barium titanate*, and *lead zirconate*. Quartz and rochelle salt are alternative materials.

The process of manufacturing ceramic transducers consists of forming the desired shape and exposing it to a high-intensity electrostatic field during the last part of a firing process.

This polarising electric field produces a mechanical strain on the molecular structure of the ceramic compound. As the ceramic cools, it gains piezoelectric properties.



Piezoelectric pressure transducer

The diagram above, shows the principle of a piezoelectric pressure transducer.

A rod connects the crystal to a diaphragm so that the crystal experiences the changing pressure applied to the diaphragm. This changing pressure causes the crystal to produce a voltage. They are not suited for static pressure measurements but are used for dynamic pressure measurements such as turbulence, ballistics and combustion.

To increase the sensitivity of this dynamic transducer, several piezoelectric crystals may be formed into a sandwich arrangement.

Level and Proximity Transducers

Level transducers

There are many ways of measuring level. However, we shall focus our attention on the following two forms of level transducers:

1. Float switch
2. Capacitive transducer.

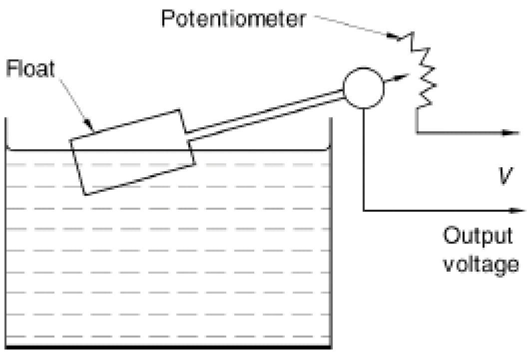
1. Float Switch System

Float switches are a common method of **measuring liquid level**. The float is either fixed or suspended in the tank, when the liquid level reaches the float, it lifts the float up and a switch inside it operates.

A very common application is in sump pumps and condensate pumps where the switch detects the rising level of liquid in the sump or tank. When the liquid reaches a pre-determined level, it energises an electrical pump which pumps the liquid out until the level has been substantially reduced. This type of control is called **on-off control**.

Another system is shown below. It is a system which uses a **potentiometer**. This method is commonly used to monitor the level in the fuel tank of a car.

The type of control is **proportional** because the actual level signal or information is made available to a control system or indicator/reader as in this case.



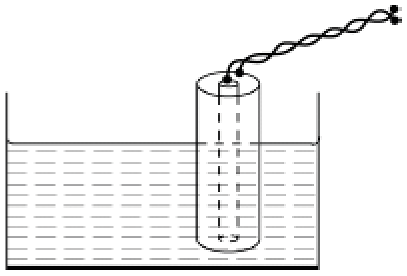
Float system using potentiometer

In the case of the fuel tank indicator, the output voltage will be proportional to the position of the wiper on the potentiometer. The resistance of the potentiometer will increase as the float moves down with a decrease in the level. This will be detected by the indicator gauge circuit.

2. Capacitive Level Transducer

Capacitive devices measure the level of solids in powdered or granular form, and liquids in vessels. They are suitable for use in extreme applications, such as measuring liquid metals (high temperatures) and liquid gases (low temperatures).

Two versions are used, according to whether the measured substance is conductive or not.



Capacitive Transducer

For non-conducting substances, two bare-metal capacitor plates, in the form of concentric cylinders, are immersed in the substance. The liquid behaves as a dielectric between the plates.

The value of capacitance measured depends on the depth of the liquid, and so level can be determined from the capacitance value.

For conducting substances, a similar method of measurement is used, but the capacitor plates are encapsulated in an insulating material.

In some cases, concentric plates may not be practical. This may be due to reasons like the likelihood of product contamination and so on. In such cases, the containing vessel may be used as one of the capacitor plates, with a single electrode inserted into the product being measured. In these instances, the instrument will need to be field calibrated to gain accuracy.

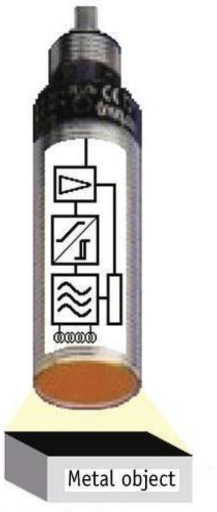
Proximity Transducer

A transducer that senses the presence of an object without making physical contact with it is called a **proximity transducer**. These devices are very useful, as they do not damage the product, and some can be mounted far away from possible damage yet still be able to monitor the flow rates of objects in many industrial processes.

There are two main categories of proximity transducers:

1. Inductive
2. Photoelectric

1. Inductive Transducer

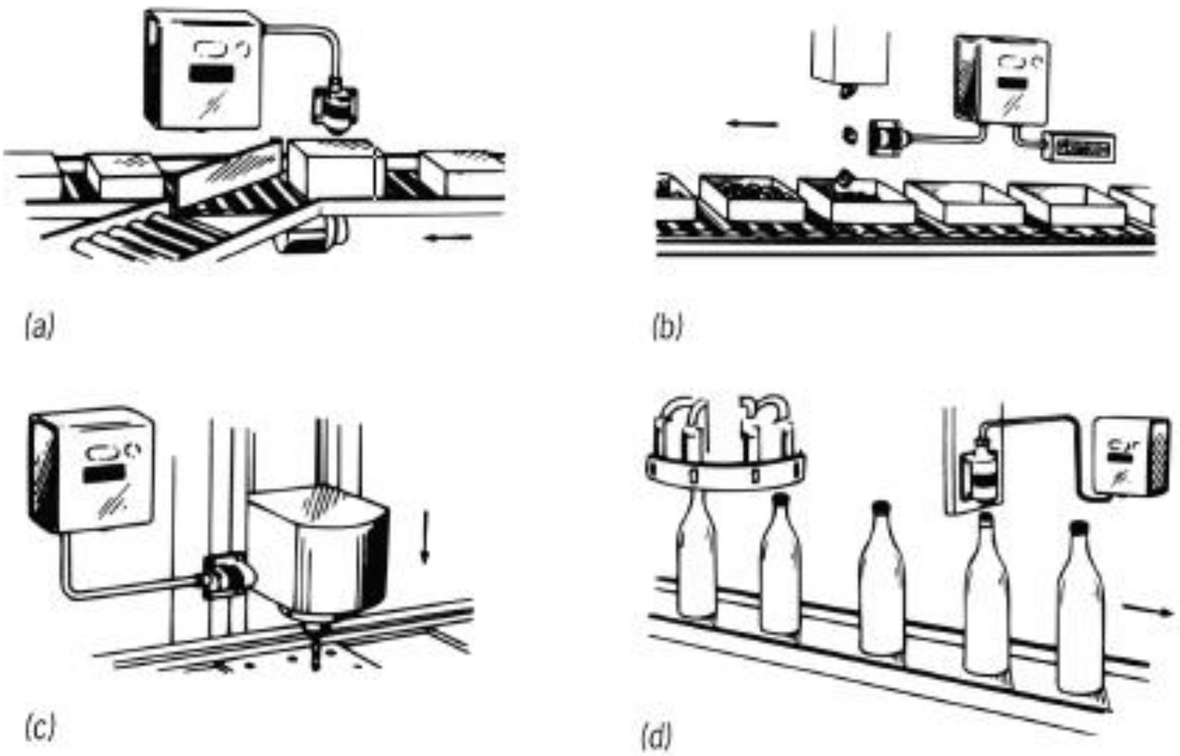


These transducers operate by detecting **changes in the magnetic field** of the area they are monitoring. These changes can be used to indicate either the presence or absence of metals (such as fragments in food products, missing tops on bottles, or objects with different size or mass). Some more advanced inductive transducers are capable of sensing and reading out the distance to object information.

The key point here is that for an inductive transducer to operate the magnetic field must be completed on the object being sensed. This means these transducers will only work on metallic objects.

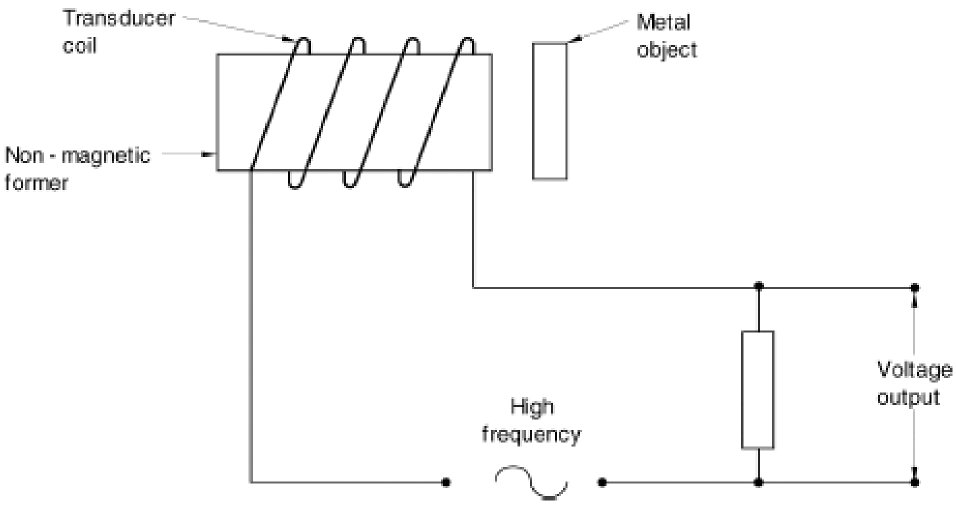
Lastly, different metals will be sensed at different distances.

The diagram below shows four possible applications of proximity transducers.



Proximity transducer applications

a) Metal boxes are sorted by height (because the proximity transducer reads the height of a box as well as its presence).
b) Steel nuts are counted as they fall past the transducer.
c) When the drill head has moved down into the sensing area of the transducer, the drilled hole is deep enough.
d) The bottle under the transducer is about to be rejected, because the transducer detects the absence of the metal cap.



Simple proximity transducer

The circuit here, shows a **simple proximity transducer**. A high-frequency a.c. supply is applied to the series circuit of the inductance coil and the resistor. The supply voltage is shared between the inductance and the resistor. In the absence of the metal object, you can assume that the voltage across the resistor is half of the supply voltage.

When the metal object is brought close to the coil, its inductance changes.

- If the metal object is made from a magnetic material, the magnetic circuit of the transducer is improved. This results in an increase in the inductance of the transducer. Inductance depends on the efficiency of the magnetic circuit and the number of turns.
- If the object is made of a non-magnetic conducting material, such as copper, the high-frequency magnetic field radiated by the transducer sets up eddy currents in the object. These currents produce a magnetic field that opposes the radiated magnetic field. This opposition reduces the effective value of transducer inductance.
- A magnetic material increases the inductance and inductive reactance of the coil, reducing the value of the output voltage. A non-magnetic conductive material has the opposite effect, reducing inductive reactance and increasing the output voltage.

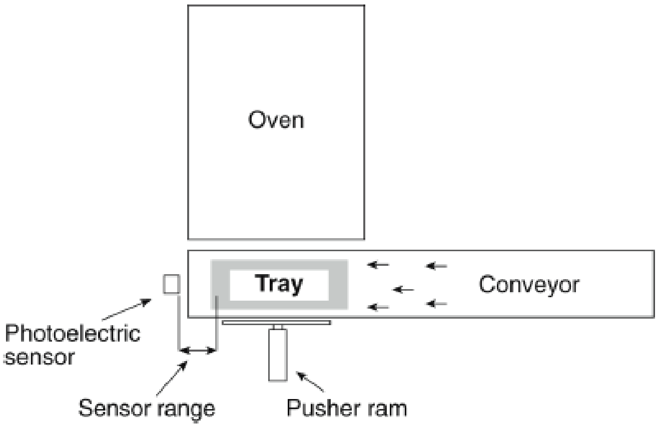
Some inductive transducers use an oscillator. The nearness of a metal object causes the inductance change to alter the frequency or stop oscillations.

Photoelectric transducers

A major advantage of photoelectric transducers over inductive transducers is that they can be mounted a long way from the material being monitored and yet remain very effective. This feature is very useful in timber plants and heavy industries, where equipment can be subject to very harsh treatment.

Photoelectric transducers are used to detect objects without contact. It's easy to think of them as a limit switch that has a beam of light instead of a mechanical arm. The sensor operates when it detects a change in the amount of light.

As you can see in the diagram (right), the sensor has detected that a tray is in position ready to be loaded into the oven.



Photoelectric sensor

The sensor is often connected to an electronic counter or controller.

There are two main types of photo-electric transducer.

1. **Beam Sensor** – This is where the light source is located in one unit, and the light receiver or photo-diode is located in another unit located some distance away.

2. **Retro-reflective type** – This is where the light source is located on one side of the transducer, and when this is reflected back from the object is read by the photo-diode sensor located on the other side of the same transducer.


Different photoelectric sensors

Transducer Output Signals

There are various standard signal ranges for process control systems and many transducers comply with these standards.

If a standard signal is not provided, then a special input interface is required in the control system. There are also a variety of convertors to change ranges if the output of a transducer does not match the input to a control system.

There are also a lot of final control elements, such as control valves, which are pneumatically controlled and as such require convertors to convert signals from electric to pneumatic.

Current signals have an advantage over voltage signals if the transducer is some distance away from the control system as there is no error caused by voltage drop in the cables.

As an output of zero can be mistaken for a broken wire, most outputs have a small signal to represent zero to allow for detection of a broken conductor.

Common current signals

The most common current signals are 0-20mA or 4-20mA.

Common Pressure Signals

There are current to pneumatic pressure convertors (I/P) to convert these common current values to 0-1bar (0-15 psi) or 0.3–1 bar (3-15 psi).

Common Voltage Signals

Voltage signals are often produced by inserting a resistor into the current loop at the receiving end. This converts the 4-20mA signal to 1-5V using a 250 Ω resistor.

There is a wider range of standard voltage ranges such as:

- 0 – 10V
- 12V
- 24V

These may also be DC or AC.

Part 2: Basic Principles of Control Systems

Two types of control systems are detailed below – these are **open loop** and **closed loop** systems. They may be manual or automatic and comprise of electronic or mechanical components.

- **Open loop** control systems rely on knowledge of the system behaviour and experience and often have a graduated control knob whose position determines the outcome without actually measuring to check if the correct outcome has been achieved. (e.g. a toaster)
- **Closed loop** systems measure the outcome, compare it with the desired outcome and adjust the control output to improve the outcome (e.g. hot water tank).

These may be manual or automatic, and their outputs **digital** or **analogue.**

A digital output is either on or off, e.g. a hot water cylinder controlled by a thermostat and an analogue output is a variable output such as depressing the accelerator on a vehicle.

Open and closed loop control systems are discussed in detail on the following pages.

Definition

Set Point (SP)	The desired value of the process variable. This could be the temperature that we want a hot water cylinder to heat up to.
Final Control Element	The device or element that controls the process. In a water cylinder this would be the heating element.
Measuring Element	This is the transducer that converts the process variable to an electrical signal. This would be a transducer converting the heat level to an equivalent electrical signal.
Feedback	The signal sent from the measuring transducer (PV) back to the comparator

Definition

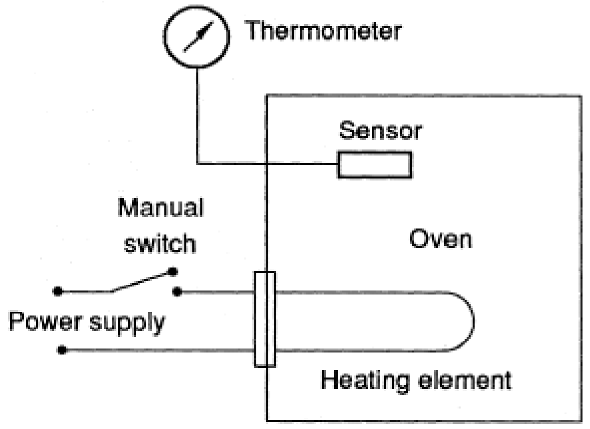
Comparator	This compares the SP with the PV and calculates the error. If the set point is a temperature of 20 degrees Celsius and the electrical signal coming into the comparator from the measuring element is equivalent to a temperature of 19 degrees Celsius, there is an error or deviation of 1 degree.
Deviation	The difference between SP and the PV
Gain	The amount that the error is multiplied by, to produce a useable output.
Controlled Variable	The output from the controller
Proportional Offset	In a proportional system, the PV will not exactly match the SP. The proportional offset is the difference between the SP and final steady state value of the PV.

Open Loop Control

This diagram shows an open loop oven temperature control.

If there are no environmental changes, a correctly designed heating element will maintain the oven at approximately the desired temperature.

As there is no way to regulate the oven temperature, this is **open loop control.** If variations of the controlled quantity (that is, the temperature of the oven) are not critical, this open loop control is quite satisfactory.



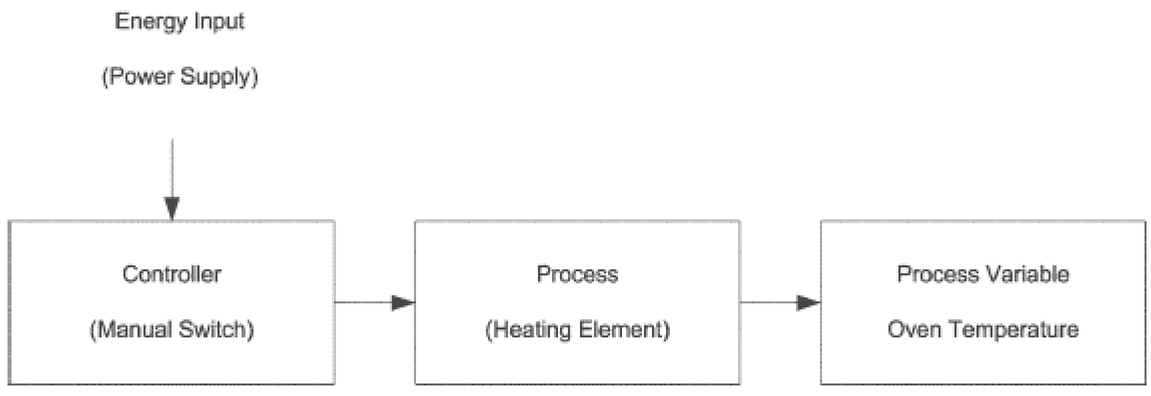
Oven temperature control

If the thermometer shows that the temperature has gone above the desired value, the operator can intervene, effectively 'closing the loop' by reading the thermometer and switching the element off. When the temperature has dropped below the desired value, the element must be switched on again.

Block diagram of open loop control

We can show each function of the process by a separate labelled block.

The diagram below shows a block diagram representing the open loop oven temperature control. The key feature to note here is that there is no automatic control or feedback from the sensor to the manual switch and in theory it would be possible for the container to overheat unless the operator intervened.



Block diagram of open loop control

In technical terms, this can be explained as follows.

The desired process value is entered into the controller. This then provides an output to the process based on the SP. The process responds to the controller and alters the output. The PV changes according to the process output. There is no feedback in this system.

Closed loop control

When the components that measure and control the oven temperature form a closed loop, it is called a closed loop control. These systems rely on feedback i.e. feeding back information on what the output is doing to the controller so an adjustment can be made if required. These systems are called **closed loop control systems** and consist of three basic functions as follows:

- Measuring
- Comparing
- Adjusting

Measuring

We cannot usefully control a parameter unless we can measure it. Every control arrangement will include a measuring device (transducer) which is able to measure the process variable (PV). In pneumatic systems, the measurement output may be in the form of air pressure; in electrical or electronic systems, it will be voltage or current signals.

Comparing

The measured value or Process Variable (PV) is compared to the desired value or Set Point Value (SP) and the deviation determined. It then adjusts the controller output or controlled variable (CV) to try and bring the two together (the CV is sometimes called the manipulated variable).

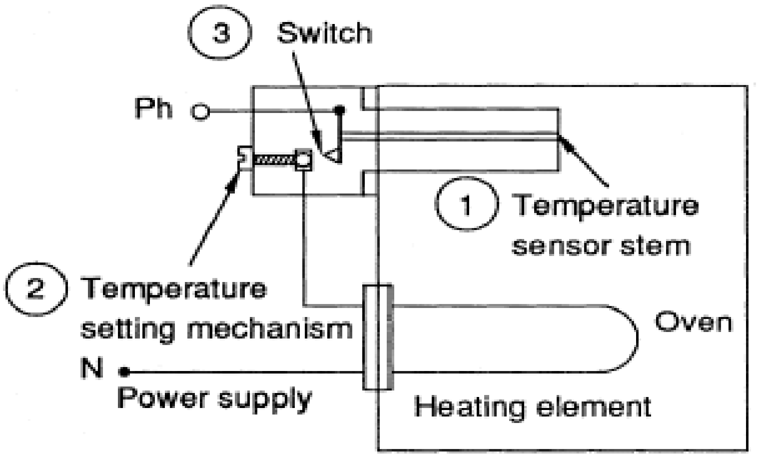
Adjusting

Once a comparison has been performed against a set point, an adjustment is required to correct this difference (or 'error') by adjusting the final control element. This brings the PV closer to the SP.

With closed loop control, the human operator can act as a part of the loop, read the thermometer, and turn the element on and off to maintain the temperature at the desired value.

However, by adding a thermostat, the human operator is no longer required to monitor and adjust the oven temperature. If the control system carries out these functions independently (it is self-regulating) or automatic.

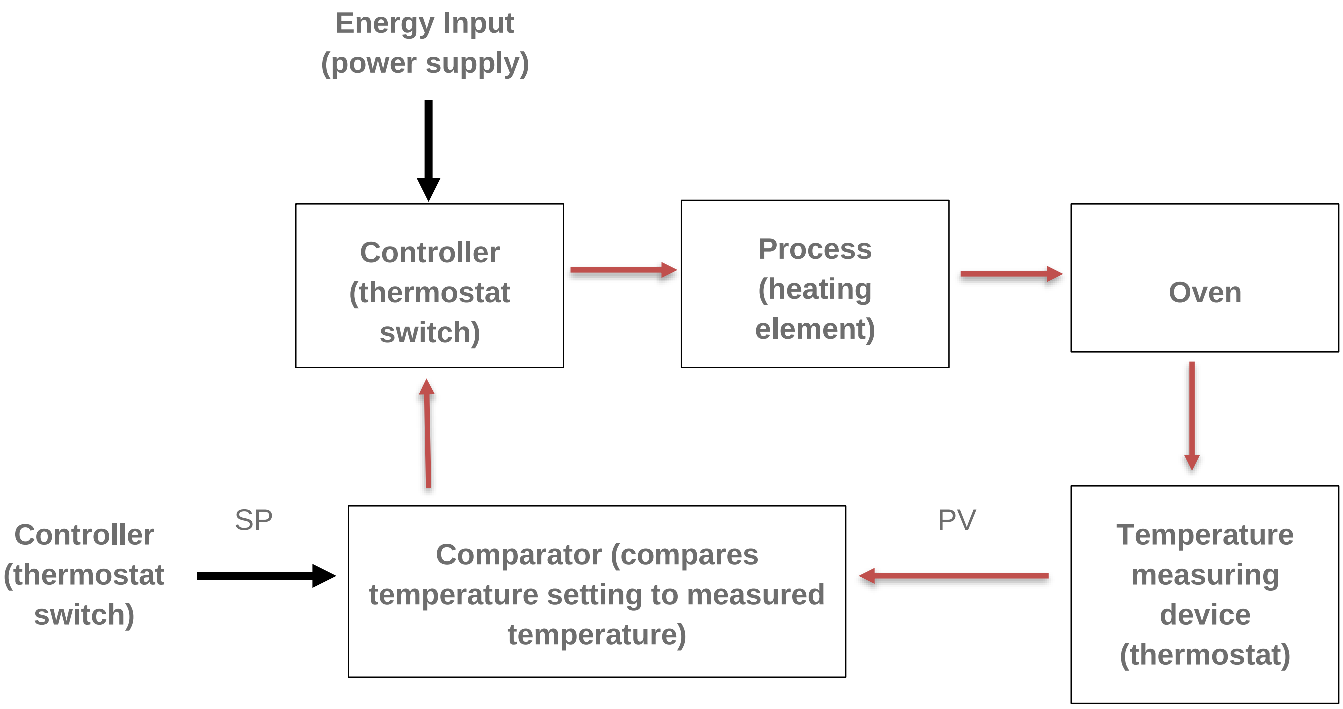
The diagram on the right, shows the measuring device (thermometer) replaced by a thermostat that allows the oven to be kept automatically at the required set temperature.



Closed loop control

The thermostat carries out **three** separate functions:

- It measures the temperature of the oven.
- It compares the temperature of the oven with the temperature setting on the thermostat.



Closed Loop Control

- It adjusts the flow of energy to the heating element by switching it on or off.

In technical terms, this can be explained as follows:

The comparator receives the measured process value (PV), and the desired value (SP), and calculates the deviation or error between these two values. The controller performs a calculation on the error and sends the result to the process or final control element. The process responds to the controller and the PV should change to a value closer to the SP. The PV changes and the result is fed back to the comparator.

Advantages of Closed Loop Control Systems

Closed loop systems have many advantages over simple open loop systems and can be used to control systems relating to:

- Levels
- Temperature
- Speed

Levels

Open Loop

The open loop system of the level of a car's fuel tank is only made into a closed loop system by the intervention of the human driver. We observe the gauge and when it gets low, we fill the tank. Without human intervention, there would be no way to 'close the loop'. The level would decrease as we used the car and in technical terms the PV would decrease and differ from the desired SP as the fuel is used.

Closed Loop

In a closed loop system, the PV would be monitored and fed into a controller to compare it to the SP. The error would be calculated, and the level would be controlled by the introduction of new liquid as it was used.

Temperature

Open Loop

We can set a heater to a specific temperature and let it heat a room. However, without feedback of the actual temperature of the room, the heater cannot react to changes in conditions of the room. Changes in the natural ambient temperature of the room would also have an effect that cannot be reacted to by an open loop system. The temperature of the room would therefore change with ambient temperature changes.

Closed Loop

In a closed loop system, the temperature of the room can be controlled closely to the set point (SP) even if the ambient temperature of the room changes. If someone opens a window or a door, the system can react to the decrease in temperature, and increase the heat applied to the room to compensate.

Motor Speed

Open Loop

An open loop control system for the speed of a motor is not able to react to changes in conditions (load changes etc) that may change its speed.

Closed Loop

In a closed loop system, the output of the motor (in the form of an electrical signal from a tacho generator) can be fed into a comparator and compared to the electrical signal from the Set Point. An error (difference in speed) can be compensated by changing the input to the motor. The motor speed can therefore be adjusted and kept close to the SP despite changes in conditions (loading, voltage etc).

Objectives of a Good Control System

The objectives of a good control system are to have a low offset to keep the PV close to the SP, a fast response to quickly respond to a changing PV, and good stability to prevent oscillations.

In terms of motor control, we want the speed of the motor (PV) to be as close to the desired speed (SP). The closer the speed, the lower the offset.

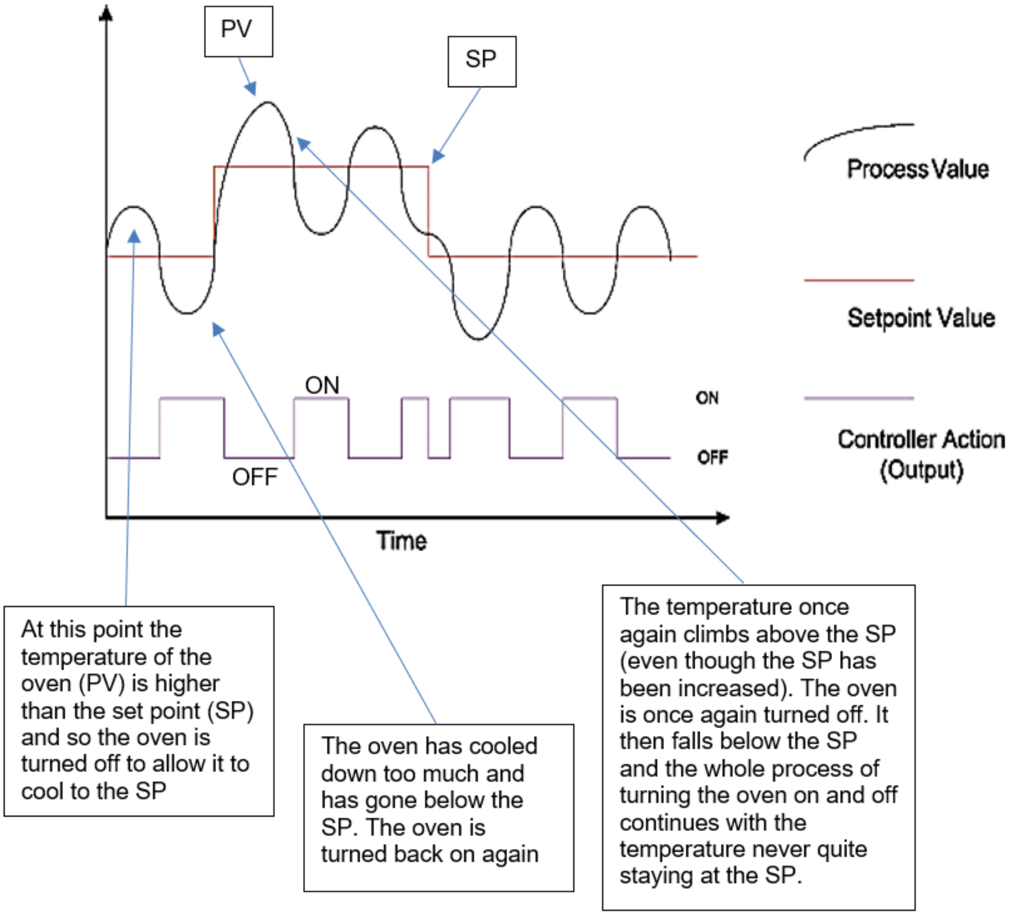
We want the controller to react quickly to the change in speed. If the motor becomes loaded and it slows down, we don't want to have a delay before the correction is made. We therefore want a fast response time.

If the controller increases the speed of the motor and then it runs too quickly, we then have to slow the motor down to get back to the SP. If the motor slows down too much, we then have to speed it back up again. This is known as oscillation. The controller is hunting for the desired speed and not quite getting it spot on each time. A good control system will keep these oscillations to a minimum.

Process response diagrams

The process response diagram below represents the reaction of the control system and the actual process we are controlling, in response to a defined change in its operating conditions.

This could be, for instance, the temperature control loop for an industrial baking oven.



The process changes can be due to several different causes; an increase in load due to the ambient temperature dropping, or additional cold items being placed in the oven.

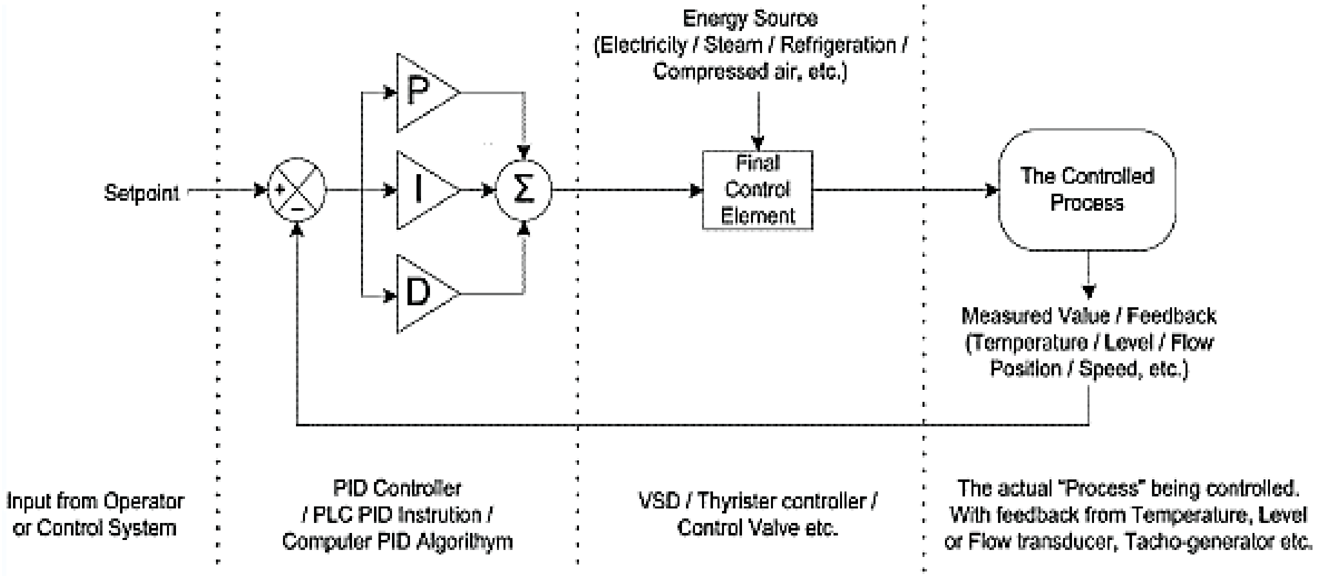
We can look at various stages of the process in the diagram below.

As is quite evident from this diagram, we need something other than just turning the controller on and off each time there is an error. That is where the three term PID controller comes in to give us a response that will keep the PV as close to the SP as possible.

Part 3: Three Term Controllers

The name “three term” comes from three types of action commonly built in to controllers. These three types of action, namely **proportional**, **integral** and **derivative** are mathematical calculations performed on the error signal and then added to derive the controller output. These three actions may be used independently or in combination by adjusting the three settings usually abbreviated P, I and D. These controllers are thus called **PID controllers**.

The next section looks at the individual P, I and D control and shows how, alone they are not able to produce very effective control but once their characteristics are combined, they can produce a near perfect closed loop control system.



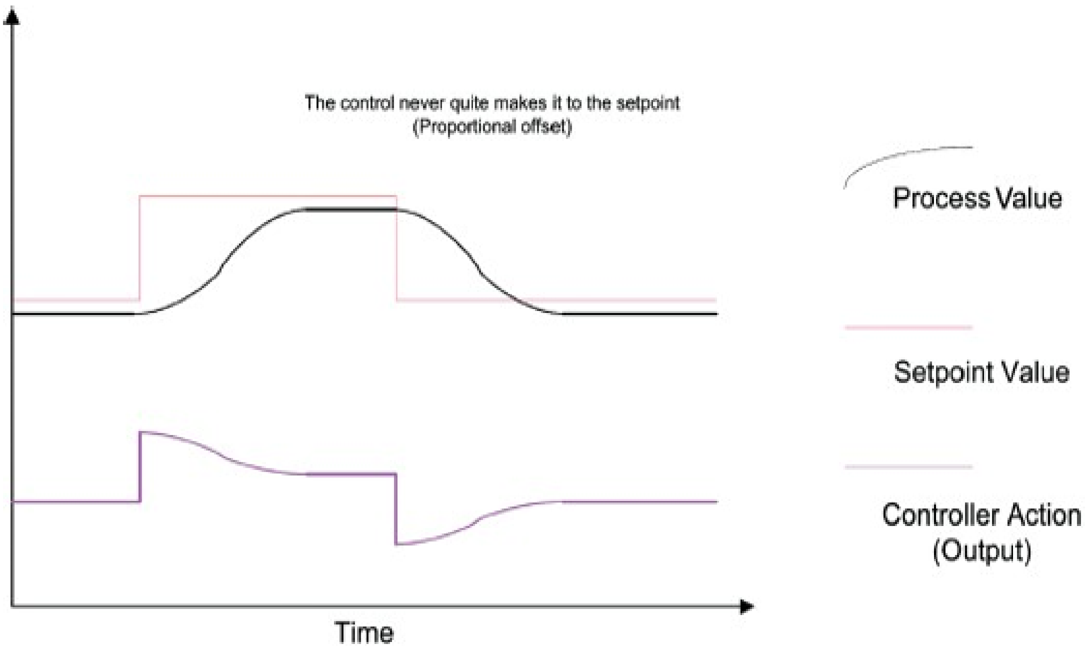
Block diagram of a ‘Parallel’ PID controller implementation showing the process relationships and components of the control loop.

Proportional Control Action

With this type of control, the controller output action is a signal which is proportional to the error between the SP and PV.

If the error is large, then the control action is large. An amplification or multiplication factor or gain can be set to determine how large the output is compared to the error. This gain is the ‘P’ setting within the controller.

Although proportional control is simple to understand, it has drawbacks. The main problem is that for most systems it will never entirely remove the error. With proportional control, the output or power to the heating element (CV) is proportional to the error. This means that the PV will never reach the SP otherwise there would be no error to drive the output.



The difference between the final steady value of the PV and the Set point value is called the **offset**.

Offset is an inherent feature of proportional control systems, and so in proportional control systems it is called **proportional offset**.

To eliminate this offset error the controller must increase its output by including another type of action called **Integral** action.

Integral Control Action

With this type of control, the output action depends on the error and **how long** that the error exists. The output increases and keeps increasing for however long the error exists.

Integrating the time is another word for this incremental or continual adding to the output while there is an error.

When the error has been reduced to zero, the adding or integrating action effectively stops. The accumulated addition is still added to the output but this value stops increasing.

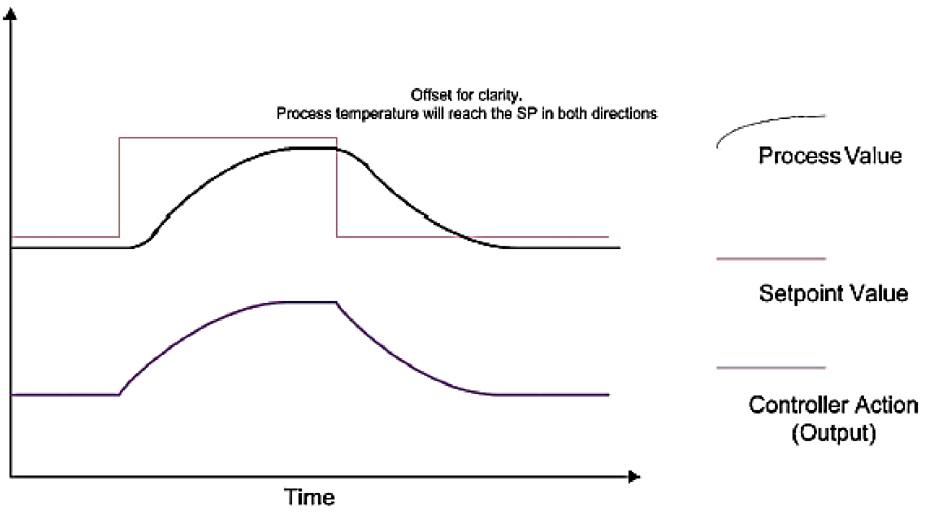
The time period intervals over which the signal is integrated or accumulated is sometimes called the reset time.

Integral-only control, while possible, is not very practical as it requires the loop to be in error for this period of time to cause a change in output, that will bring the process to its new set point.

This causes the control loop to be sluggish and show a process lag or time delay in its action.

When combined with proportional control (PI), it will compensate for the proportional offset (steady state error) and allow the loop to exactly reach the set point.

This combination (PI control) is by far the most commonly used analogue control loop in continuous process industries.



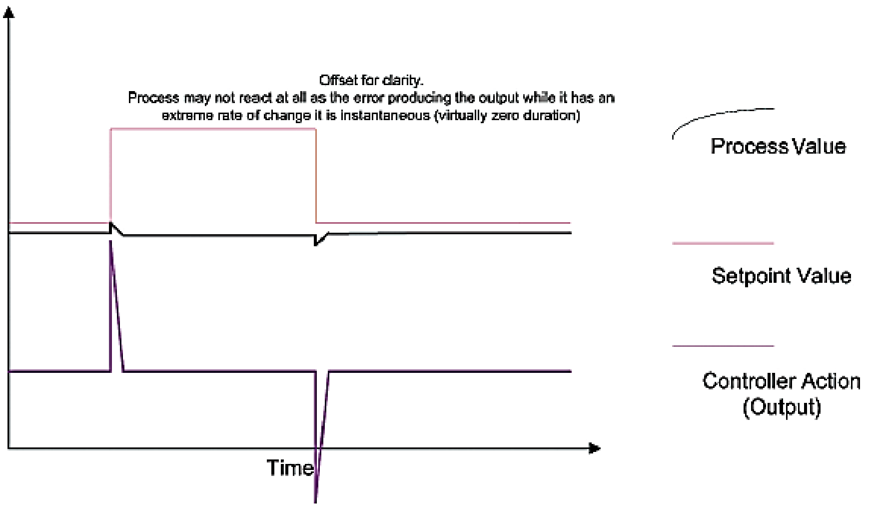
Definition

Integral control action	Control, in which the degree of correcting action is determined by the amount of error we have accumulated over a specific time period (the reset time).

Derivative (or Rate) Control Action

With this type of control, the output action depends on the error and **how fast** it is changing.

By monitoring the speed, or derivative, of the error it anticipates how much more or how much less output is required to eliminate the error in an optimal time.



Definition

Derivative control action	Control in which the degree of correcting action is determined by the rate at which the error producing it changes.

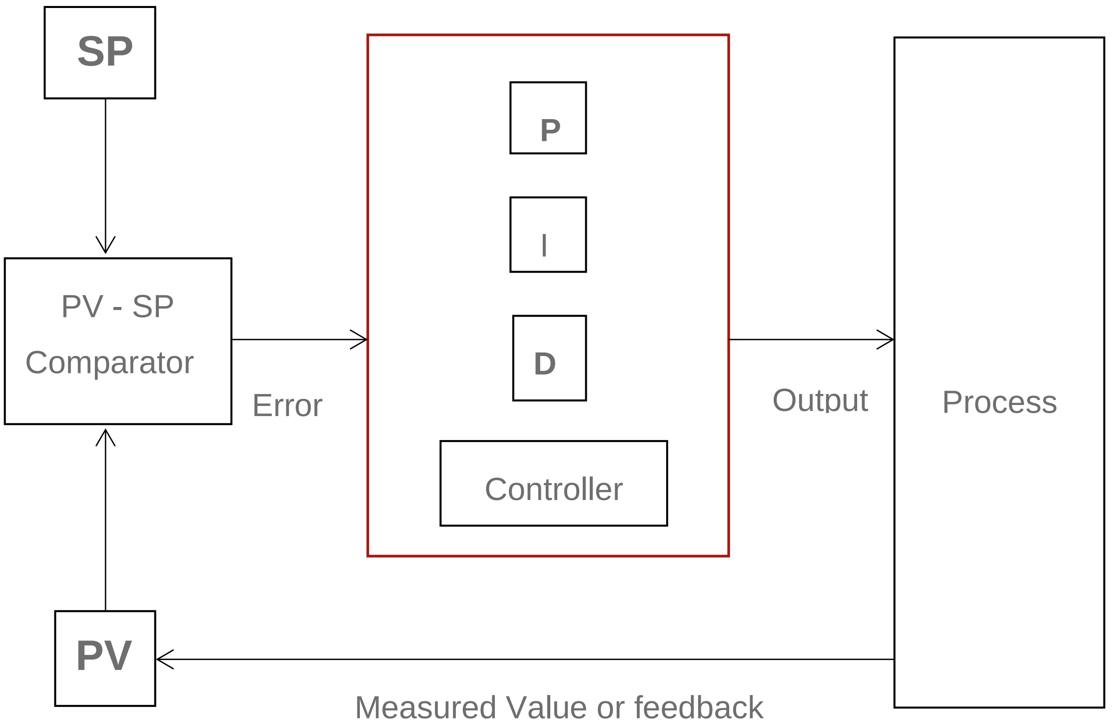
Derivative-only control, while possible, is usually impractical, as it reacts only to a changing error.

If this control is applied to the process input rather than the error, the output can be made to resist change, as in force feedback joysticks used by electronic games. The faster you try to move the joystick, the more it resists your action.

Derivative action is typically used with P or PI to compensate for rapidly changing input values from the process feedback.

The PID controller

The diagram below shows a PID controller. It combines the characteristics and advantages of the individual P, I, and D controllers to produce a control that will keep the PV as close as possible to the SP with minimum error, low offset, good response times and good stability. The tuning of the PID components is crucial to achieving this and will be different for different applications.



Block diagrams of Closed Loop PID Control

Ideally the combination of the three control actions should cause the PV to reach the SP as quickly as possible. If the actions are too fast, then overshoot and undershoot occur where the PV reaches the SP but the action continues and the SP is passed.

Overshoot and undershoot may also continue and even increase causing the control system to become unstable.

Tuning a PID controller is **outside the scope of this unit standard** as it is a demanding and complex problem and best left to a controls engineer.

Summary of PID Control:

The PID controller is used to control a process using a combination of the three controllers. It combines proportional, integral and derivative controllers to provide the advantages of each in one package.

An example would be the control of the speed of a motor.

- The **SP** (set point) would be the desired speed of the motor.
- The **process** would be the motor.
- The **process variable** is the speed in an electrical form that could be produced by a tacho generator connected to the shaft of the motor.
- The **signal** produced by the tacho generator would be proportional to the speed output of the motor.

The PV to SP comparator would look at the signals from the SP and the PV and look for a difference.

- If there is **no difference**, (no error) the motor is turning at the desired speed and there is no error signal produced.
- If there is **a difference**, an error signal is sent to the PID Controller. It uses a combination of PID control to produce an output that it will send to the motor to either increase or decrease its speed. The change in speed will bring it closer to the desired speed, which will therefore decrease the error, which in turn will decrease the output to the motor. This process continues until the desired speed is met.

An increase in loading of the motor could decrease its speed and the process would start again.

Te Pūkenga

earnlearn-tepukenga.ac.nz
0800 327 648 (0800 EARN IT)